A mathematical model of serious and minor criminal activity

Citation for published version:
Lacey, AA & Tsardakas, M 2016, 'A mathematical model of serious and minor criminal activity', *European Journal of Applied Mathematics*, vol. 27, no. 3, pp. 403-421. https://doi.org/10.1017/S0956792516000139

Digital Object Identifier (DOI):
10.1017/S0956792516000139

Link:
Link to publication record in Heriot-Watt Research Portal

Document Version:
Publisher's PDF, also known as Version of record

Published In:
European Journal of Applied Mathematics

European Journal of Applied Mathematics

http://journals.cambridge.org/EJM

A mathematical model of serious and minor criminal activity

A. A. LACEY and M. N. TSARDAKAS

European Journal of Applied Mathematics / Volume 27 / Special Issue 03 / June 2016, pp 403 - 421
DOI: 10.1017/S0956792516000139, Published online: 08 April 2016

Link to this article: http://journals.cambridge.org/abstract_S0956792516000139

How to cite this article:
A. A. LACEY and M. N. TSARDAKAS (2016). A mathematical model of serious and minor criminal activity. European Journal of Applied Mathematics, 27, pp 403-421 doi:10.1017/S0956792516000139

Euro. Jnl of Applied Mathematics (2016), *vol.* 27, *pp.* 403–421. © Cambridge University Press 2016

doi:10.1017/S0956792516000139

A mathematical model of serious and minor criminal activity

A. A. LACEY and M. N. TSARDAKAS

Department of Mathematics, Heriot-Watt University, Edinburgh, EH14 4AS, UK
email: a.a.Lacey@hw.ac.uk, mtsardakas@gmail.com

(*Received 27 April 2015; revised 29 February 2016; accepted 1 March 2016; first published online 8 April 2016*)

Using mathematical methods to understand and model crime is a recent idea that has drawn considerable attention from researchers during the last five years. From the plethora of models that have been proposed, perhaps the most successful one has been a diffusion-type differential equations model that describes how the number of criminals evolves in a specific area. We propose a more detailed form of this model that allows for two distinct criminal types associated with major and minor crime. Additionally, we examine a stochastic variant of the model that represents more realistically the 'generation' of new criminals. Numerical solutions from both models are presented and compared with actual crime data for the Greater Manchester area. Agreement between simulations and actual data is satisfactory. A preliminary statistical analysis of the data also supports the model's potential to describe crime.

Key words: Crime, differential equations, mathematical model

1 Introduction

Crime has always been an inherent part of every human society. It is not surprising that researchers expend a considerable amount of time and resources in an attempt to understand the way crime is generated and shaped [2]. Mathematical modelling is a newcomer in the fight against crime, with a number of models being proposed during the last five years that use a variety of approaches to model crime. These vary from agent-based models [5, 13], population dynamics models [16] and epidemiological models [23] to game-theoretic [14, 19] and probabilistic models [7, 15, 22].

Lately, a significant amount of research has been conducted on the spatial and temporal distribution of crime [3, 6]. Specifically, the formation of areas of increased criminal activity (hotspots) and their evolution in time and space has drawn the attention of crime scientists [9, 10]. A novel mathematical modelling approach in attacking this problem is presented in a series of papers by Short *et al.* [18, 20, 21], where the authors derive a continuous model that captures the behaviour of burglars and which leads to the formation of hotspots.

A common denominator in most work to date is the type of crime the research deals with. More often than not, researchers either use 'crime' as a generic term to describe all

Table 1. *Categorisation of criminal activities present in the data*

Serious	Minor	Discarded
Burglary	Public disorder & weapons	Drugs
Robbery	Shoplifting	Other crime
Criminal damage & arson	Vehicle crime	Anti-social behaviour
Violent crime	Other theft	
Violence & sexual offences	Theft from person	
	Public order	
	Possession of weapons	
	Bicycle theft	

crime which can occur or they tend to focus to a specific crime category, e.g. burglaries. We believe that a visualisation of the way 'serious' crime (such as burglaries or violent crime) and 'minor' crime (e.g. shoplifting) evolve and interact would be interesting and could aid researchers towards a more thorough understanding of crime.

The goal of this paper is twofold. First, we aim to explore the relation between serious and minor crime in an area from a data-analysis viewpoint. Second, we propose a quantitative mathematical model that captures the dynamics that govern the temporal distribution of serious and minor crime, as identified through the analysis of the data. We make no attempt to base our model on social or psychological theories of crime, electing instead to focus on it being consistent with the results from the data analysis. We consider both deterministic and stochastic versions of the model and exhibit its suitability in describing crime evolution via comparison with the actual data. Finally, parameter values for the model are extracted from the data using a least squares fit.

2 Data analysis

Examining data of criminal activities of more than one type could be illuminating and help set the basis for our model. To this end, we use a dataset which contains all reported crimes in England for a period of two and a half years, obtained from the UK police website [4]. Each criminal activity is categorised into one of 16 different types. Apart from this categorisation, the month each crime is committed as well as an approximate location is available. Despite the uncertainty present in the dataset (monthly basis instead of daily, imprecise locations), we may safely use it for an initial macroscopic analysis of crime patterns, as evidenced in [24].

As a first approach, let us examine the distribution of crime in time in some metropolitan areas. We therefore merge all datasets for each area and aggregate the number of crimes of each type on a monthly basis. Because of the large number of categories available, we combine some of these into two larger, arbitrarily defined categories, serious and minor crime, to simplify the analysis. Discarding some crime types is also useful, due to their ambivalent nature. Communication with crime scientists indeed confirmed that labels such as 'drugs' tend to be unclear and should best be omitted within the context of an initial analysis. The detailed categorisation is shown in Table 1. After the data have been cleaned and formatted, we can plot time series which describe the evolution of the total

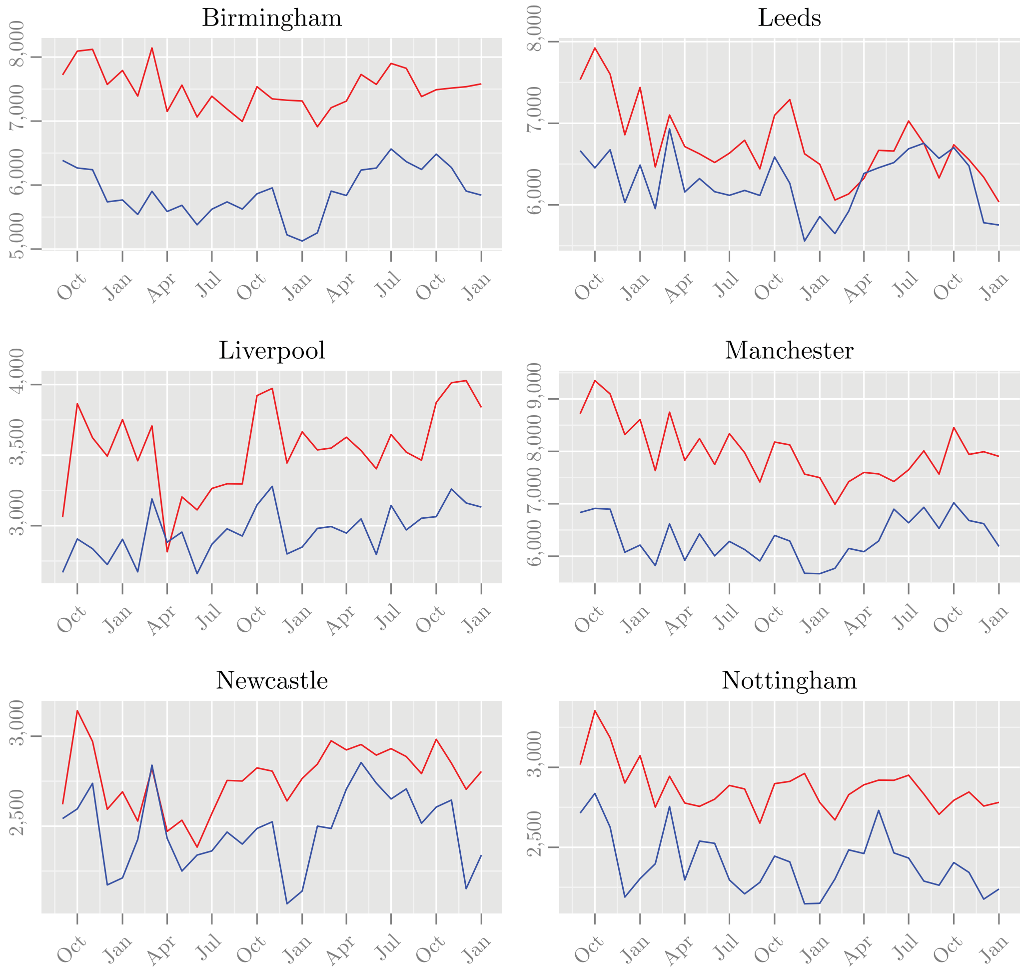

FIGURE 1. Total serious (red) and minor (blue) criminal activity per month between September 2011 and January 2014.

number of crimes (serious and minor) in a specific area on a monthly basis, as shown in Figure 1.

The first impression one gets from studying the time series is the way serious crime mimics the behaviour of minor crime (and vice versa). Local maxima (resp. minima) of serious crime coincide with those of minor crime in the overwhelming majority of cases. This trend is very strong in all metropolitan areas examined and is something that should definitely be taken into account when modelling this situation. A further trend that is evident in a number of cities, albeit to a different degree, is that of periodicity on a yearly basis. This phenomenon is not nearly as strong as the first trend we noticed and needs careful statistical testing to confirm. However, the literature on crime and seasonality generally suggests that this is usually the case (see [8] for seasonality in property theft and [1] for more crime categories).

Table 2. *Pearson's r calculated using data from September 2011 to January 2014. The coefficients are in the left columns whereas the respective p-values are on the right. S and M represent serious and minor crime respectively and ΔS and ΔM are forward differences, e.g.* ΔS_{June} *= (Total amount of serious crime in July) − (Total amount of serious crime in June)*

	Pearson's r											
	$S-\Delta S$		$M-\Delta M$		$S-M$		$\Delta S-\Delta M$		$S-\Delta M$		$M-\Delta S$	
Birmingham	−0.58	0.01	−0.48	0.02	0.67	<0.01	0.58	<0.01	−0.56	0.01	−0.09	0.69
Manchester	−0.52	<0.01	−0.52	<0.01	0.58	<0.01	0.80	<0.01	−0.58	<0.01	−0.29	0.13
Leeds–Bradford	−0.47	0.01	−0.55	<0.01	0.58	<0.01	0.71	<0.01	−0.46	0.01	−0.22	0.26
Liverpool–Merseyside	−0.58	<0.01	−0.66	<0.01	0.67	<0.01	0.71	<0.01	−0.35	0.07	−0.55	<0.01
Newcastle–Sunderland	−0.54	<0.01	−0.55	<0.01	0.61	<0.01	0.56	<0.01	−0.22	0.26	−0.38	0.04
Nottingham–Derby	−0.53	<0.01	−0.57	<0.01	0.57	<0.01	0.45	0.02	−0.51	0.01	0.01	0.98
Entire dataset	−0.61	<0.01	−0.60	<0.01	0.69	<0.01	0.87	<0.01	−0.65	<0.01	−0.43	0.02

To further test the observation that both crime types have the same qualitative behaviour, we can calculate correlation coefficients in our sample. We mainly use Pearson's r, a coefficient ranging from -1 to 1 which shows trends of linear relationship between two variables. A value of 1 or -1 indicates that the relationship between two variables is perfectly described by a linear function, whereas smaller (in absolute value) coefficients indicate lower correlation. Additionally, we calculate Spearman's ρ, another measure of statistical dependence that ranges from -1 to 1, which assesses how well the relationship between two variables can be described using a monotonic function. The statistical significance of both correlations can be tested using the Fisher transformation and calculating the relevant p-values. A p-value lower than a certain threshold (here 0.05) signifies that we can reject the null hypothesis of the variables being independent. Some results of the calculations can be seen in Table 2, whereas the full results can be seen in Table A 1. S and M denote serious and minor crime respectively, whereas ΔS and ΔM are forward differences such that if S_i is the total crime rate for month i, we have $\Delta S_i = S_{i+1} - S_i$. One can think of these forward differences as approximate discrete derivatives that provide a way of measuring the rate of change of S and M. Hence, correlations between them and S and M can indicate the functional dependence of these derivatives on S and M.

Each half of the first two columns in Table 2 indicates that there exists a strong negative correlation between S and ΔS (resp. M and ΔM). Hence, when serious crime S increases, its rate of change ΔS will tend to decrease whereas when S decreases, ΔS will increase. Therefore, a negative feedback loop is formed for these variables, which will revert S to its mean value whenever it jumps significantly. This kind of behaviour can be explained by having criminals being removed from the system more rapidly as the crime rate increases; more crime leads to more arrests and sentencing [11]. The same reasoning can explain the relation between M and ΔM as well.

The third and fourth columns in Table 2 reinforce our initial impressions from the time series as it shows the existence of a strong linear correlation between serious and minor crime, with a Pearson's r of 0.69 for the aggregated data. ΔS and ΔM exhibit the same kind of relation, indicating that the serious and minor crime rate vary in the same way.

Finally, we have calculated the correlations between S and ΔM and between M and ΔS. These computations aim to see if there is any correlation between one crime type and the rate of change of the other, which seems to be the case but only for the $S{-}\Delta M$ coefficient, as most coefficients for the other case are not statistically significant. This suggests that serious crime in an area will tend to influence minor crime, a phenomenon which is frequently observed in areas where organised crime is present. For instance, members of gangs may 'control' the amount of minor crime taking place in areas where they are active. We shall propose, in the following section, a mechanism by which one crime type can have effect, albeit indirect, on the other.

At this point, it should be mentioned that this analysis was also carried out for a number of rural areas as well. While a full presentation and comparison with the metropolitan areas analysis is outside the scope of this paper, it is important to state that there are no significant differences in the temporal distribution of crime between urban and rural areas (see Table A 1). A calculation of the correlation coefficients yields similar results to those of the metropolitan areas, although the significant correlations tend to have a slightly lower absolute value. At the moment, it is unclear whether this is a significant differentiation which would lead to a different model. Initially, we will focus on the findings for the metropolitan areas and we hope to carry out a more extensive comparison between the two types of areas in the future.

Summing up this preliminary data analysis, there are two main points that need to be kept in mind while constructing the model. First, serious crime and minor crime should exhibit the same qualitative behaviour and second, both crime categories should form a negative feedback loop with their rate of change/time derivative. We will now propose a continuous model that presents these properties.

3 Deterministic model

3.1 Overview

Our main goal is to obtain a model that describes the evolution of crime in a certain area and exhibits behaviour close to the one shown in Figure 1 and with the properties of Table 2. To this end, let us examine a model that consists of two different types of criminals, serious and minor. We will use $\rho_1(t)$ (resp. $\rho_2(t)$) to denote the number of serious (resp. minor) criminals active in our area at time t. In the spirit of Short *et al.* [21], the behaviour of the criminals is driven by a quantity which we will refer to as the *attractiveness* of the area. One can think of the attractiveness as an indicator of how probable it is for a criminal to act at a specific time. To increase the flexibility of the model, we allow the attractiveness to depend not only on the behaviour of the active criminals but also on other factors such as time, characteristics of the area examined, or the type of crime committed. With this in mind, we split the attractiveness as follows:

$$\text{Attractiveness} = A(t) + B(t).$$

Here, $A(t)$ denotes the 'intrinsic' part of the attractiveness that depends on factors other than the behaviour of criminals and $B(t)$ represents the 'dynamic' part of the attractiveness that is caused by criminal activity. To be more concrete, let us suppose that knowledge of crimes being committed in an area tends to encourage more crimes to take place. This effect would then be represented by the dynamic $B(t)$ term. Conversely, if the number of police officers patrolling a certain area changes according to the number of crimes taking place, that would be a negative effect represented again by $B(t)$. On the other hand, changes in attractiveness due to factors not affected by criminal activity (e.g. time of day or seasonality) will be accounted for by the intrinsic attractiveness $A(t)$.

We will now discuss the behaviour of the criminals ρ_1 and ρ_2. Let us assume that at a certain time t a number of individuals commit a crime. Some of those are arrested and therefore removed from the system, whereas others appear in the system, perhaps due to release from prison or through people becoming criminals. We first consider how the number of criminals evolve. We take the rate of lost criminals, through arrest and conviction, to be a constant multiple of the rate at which crimes are committed, namely $k_i\rho_i\,(A+B)$, $i=1,2$. Because of the way attractiveness is defined, we assume that the total number of crimes of type i committed at time t is proportional to the product of the total attractiveness by the number of criminals, resulting in a contribution to the rate of change of the form

$$-k_i c_i \rho_i\,(A(t)+B(t))\,, \quad i=1,2,$$

where each k_i and c_i are constants of proportionality. It is however considerably more difficult to propose a form for the term which represents the generation of new criminals. To simplify the analysis to follow, we will consider it as constant and name it γ_1 (γ_2) for the creation of serious (resp. minor) criminals. Therefore, the evolution in time of the number of criminals of either type is described by the following equation:

$$\frac{d\rho_i}{dt} = \text{replacementcriminals} - \text{constant} \times \text{criminals who are committing a crime}, \quad i=1,2.$$

These can be rewritten in more detail as two evolution equations for our two types of criminals:

$$\frac{d\rho_1}{dt} = \gamma_1 - k_1 c_1 \rho_1\,(A+B)\,, \tag{3.1}$$

$$\frac{d\rho_2}{dt} = \gamma_2 - k_2 c_2 \rho_2\,(A+B)\,. \tag{3.2}$$

Let us now examine the behaviour of the dynamic part of the attractiveness, $B(t)$. Every crime that is committed increases $B(t)$ and therefore the dynamic attractiveness is boosted by a term proportional to the total number of crimes of both categories committed. We use the term

$$(\lambda_1\rho_1 + \lambda_2\rho_2)\,(A+B)$$

to model this boost, where λ_1 and λ_2 are constants. Note that we have implicitly assumed that the dynamic attractiveness $B(t)$ is global rather than local, in the sense that criminals may exchange information about crimes committed. This is consistent with the growth part of the attractiveness in the spatio-temporal model in [21]. We further assume that B decays exponentially in time. Hence, the evolution equation for this part of the attractiveness is

$$\frac{dB}{dt} = (\lambda_1\rho_1 + \lambda_2\rho_2)\,(A+B) - \omega B, \tag{3.3}$$

where ω is the (constant) decay rate. This equation, together with equations (3.1) and (3.2), forms a 3×3 non-linear coupled system of ODEs. Although this model form is mathematically similar to the one of Short *et al.* [21], the interpretations of the two models vary considerably. Whereas the Short model takes into account diffusion of both criminals and attractiveness as well as local spatial effects, the model proposed above describes the temporal evolution of two types of criminals in a larger scale.

We can now confirm the existence of a negative feedback between both crime types and their derivatives, a characteristic we noted in the data analysis section as well. When serious crime S increases, the $(\lambda_1\rho_1 + \lambda_2\rho_2)(A + B)$ term in equation (3.3) will increase as well. This will boost the dynamic part of the attractiveness B which will in turn cause the $-k_1c_1\rho_1(A + B)$ term in equation (3.1) to decrease, ultimately decreasing $d\rho_1/dt$. The same hold for ρ_2 and its derivative as well. At this point, we should note that the behaviour of ρ_1 and ρ_2 is, in general, symmetric, mirroring the behaviour of serious and minor crime in the data analysis. Despite the symmetric form of all three equations, it is possible to introduce asymmetry in the model in an indirect way by altering the sizes of λ_1 and λ_2.

An alternative form of the model can be obtained if we assume that the dynamic attractiveness B is more 'personal', that is information is not shared between criminals so that only those who have committed a crime are aware it has taken place. This would replace the $(\lambda_1\rho_1 + \lambda_2\rho_2)(A + B)$ term in equation (3.3) with a term of the form $c(A + B)$, resulting in an uncoupled equation for the attractiveness. This simpler version however does not link ρ_1 and ρ_2 and is consequently not consistent with the results of the data analysis. Therefore, we will continue our analysis with the more complicated form of the model because of its greater flexibility and therefore increased ability to fully grasp the dynamics at play.

3.2 Model analysis

As a first step, we can write the equations of the system in non-dimensional form. To simplify the analysis to follow, let us define $K_i = c_ik_i$, $i = 1, 2$. We begin by looking for natural scales of quantities present in the model. A natural choice for the time scale is $\tau = 1/\omega$, as this is the relaxation time for the dynamic attractiveness B. We also scale ρ_1 (resp. ρ_2) by ω/λ_1 (resp. ω/λ_2) as this gives a balance of the B terms on the right-hand side of equation (3.3). Finally, B is scaled in a way to balance the left-hand side (rate of change of serious crime) and the second term (criminal removal rate) on the right-hand side of equation (3.1). We therefore scale variables as follows, denoting dimensionless quantities with a hat:

$$\hat{t} = \frac{1}{\omega}t, \quad \hat{\rho}_i = \frac{\omega}{\lambda_i}\rho_i, \quad \hat{B} = \frac{\omega}{K_1}B.$$

At the moment, we have limited information on the magnitude of the intrinsic part of the attractiveness A and will therefore scale it in the same way as the dynamic attractiveness:

$$\hat{A} = \frac{\omega}{K_1}A.$$

Finally, we introduce the dimensionless constants

$$\hat{\gamma}_i = \frac{\omega^2}{\lambda_i}\gamma_i, \quad K = \frac{K_2}{K_1}, \quad i = 1, 2$$

and after dropping the hat notation, the system of equations can now be written as

$$\frac{d\rho_1}{dt} = \gamma_1 - \rho_1 (A + B),$$
$$\frac{d\rho_2}{dt} = \gamma_2 - K\rho_2 (A + B),$$
$$\frac{dB}{dt} = (\rho_1 + \rho_2)(A + B) - B.$$

For simplicity, let us now consider the above system with the intrinsic attractiveness being constant, $A(t) = A$. It is easy to see that there exists a unique equilibrium point, namely

$$\bar{\rho}_1 = \frac{K\gamma_1}{AK + K\gamma_1 + \gamma_2},$$
$$\bar{\rho}_2 = \frac{\gamma_2}{AK + K\gamma_1 + \gamma_2},$$
$$\bar{B} = \frac{K\gamma_1 + \gamma_2}{K}.$$

We can now use the Routh–Hourwitz criterion to prove that the steady state is stable for any choice of parameter values. One of the necessary and sufficient conditions for stability of the steady state, namely that $K\,(A + B)^2 > 0$, is automatically satisfied as K is positive. By substituting the equilibrium values for ρ_1, ρ_2 and B, the second condition can be written as

$$(K + 1)(\gamma_1 + \gamma_2) + A\left(\frac{K}{K(A + \gamma_1) + \gamma_2} + K + 1\right) > 0,$$

which obviously holds as well. Similarly, we can prove that the final condition of the Routh–Hourwitz criterion holds as well and therefore the steady state is always stable.

4 Numerical results

Despite the innocuous form of the equations of the model, it may be impossible to find an exact solution. Thus, we resort to numerical simulations to gain some insight into the behaviour of the model. The differential equations are integrated using MATLAB's ODE15 integrator. The reasoning behind choosing a stiff integrator can be justified by our lack of information regarding the parameter values and the initial data. For the simulations presented here, we use a sample collection of parameters that are similar to those used in [21]. These do not necessarily represent realistic values, as at this point our main goal is to examine the form of the solutions rather than a model that accurately describes criminal activity quantitatively.

Example output from the simulations can be seen in Figure 2. Serious crime $\rho_1(t)$is represented by a blue line, minor crime $\rho_2(t)$ by an orange line and dynamic attractiveness $B(t)$ by a green line, whereas t denotes time in days. All simulations were run with initial data away from equilibrium when A is constant, namely $\rho_1(0) = \bar{\rho}_1 + 0.5$, $\rho_2(0) = \bar{\rho}_2 + 0.5$. When $A(t)$ is periodic, all initial conditions are equal to $\left(\bar{\rho}_1, \bar{\rho}_2, \bar{B}\right)$, where we have taken

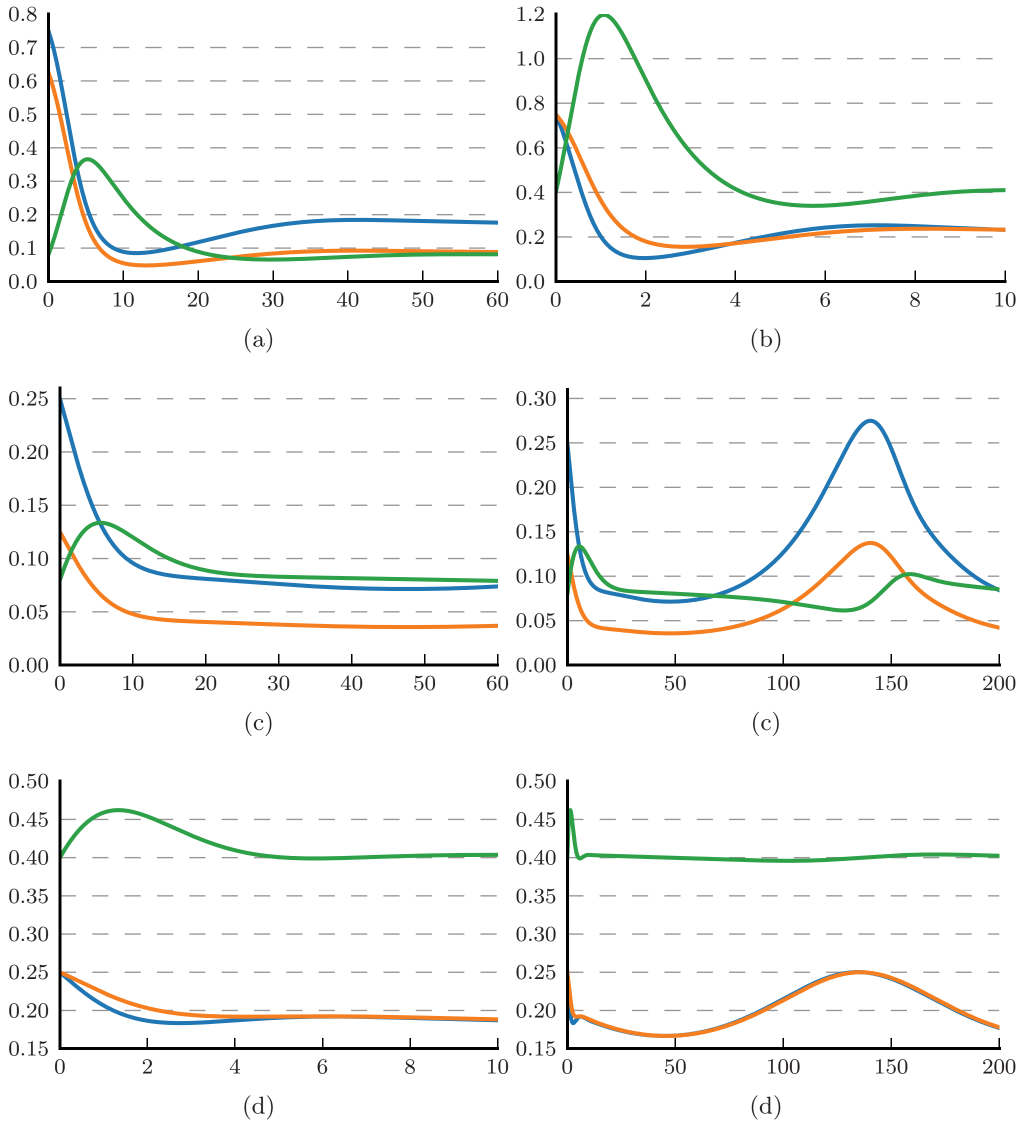

FIGURE 2. Numerical simulations using parameters described in Table 3. The x-axis represents time in days, the blue line represents serious crime ρ_1, the orange line represents minor crime ρ_2 and the green line is the dynamic attractiveness B. When A is constant, the initial condition for ρ_1 (resp. ρ_2) is equal to $\bar{\rho}_1 + 0.5$ (resp. $\bar{\rho}_2 + 0.5$). When A is periodic, the initial conditions are at equilibrium. Initial attractiveness B is always at equilibrium $\bar{B}$.

the periodic part of $A(t)$ to be equal to zero at the equilibrium formulas. The rest of the parameters for these plots are as described in Table 3.

We observe that when the intrinsic attractiveness $A(t)$ is constant (plots a and b), all variables settle down to their equilibrium values after some initial variation. On the other hand, when the intrinsic attractiveness is periodic, it is obvious that the system is driven by this periodicity (plots c and d). The short-term behaviour of the system does

Table 3. *Parameter values used for numerical simulations*

γ_1	γ_2	K_1	K_2	λ_1	λ_2	ω	$A(t)$	Plot
0.02	0.01	1	1	0.7	0.2	0.2	$1/30$	a
0.2	0.1	2	1	2.5	1.5	1	$1/30$	b
0.02	0.01	1	1	0.7	0.2	0.2	$\left(1 + \sin \frac{\pi t}{6}\right)/10$	c
0.2	0.1	2	1	2.5	1.5	1	$\left(1 + \sin \frac{\pi t}{6}\right)/10$	d

not change significantly when moving from constant A to periodic A, as is evident by comparing plots (c) and (d). The resulting long-term behaviour is close to the way the number of criminals varies in areas as in Figure 1 and therefore this type of model should be considered more realistic than the former. It should be noted that in both models, the behaviour of the two criminal types is very similar, which was one of the properties that was strongly evident during the data analysis. At this point, it is clear that the qualitative behaviour of the model is not far from that of the criminal time series, given that we are using a deterministic model to describe a process which is inherently random in nature. In an effort to further increase the flexibility of the model, we will now examine a stochastic version of it.

4.1 Stochastic model

Let us begin by identifying possible sources of randomness in our model. Some obvious candidates for that role would be the intrinsic attractiveness A (due to random fluctuations in the attractiveness of the area examined) or the coefficients K_i (random fluctuations in the effectiveness of enforcement agencies). For the sake of simplicity, we will begin by incorporating the randomness of the criminal generation rates, γ_1 and γ_2. Indeed, from a criminological point of view, there is no reason why the 'generation' of new criminals should not be a random process. To model this, let us rewrite the equations for the number of criminals with the criminal generation rates being stochastic instead of constant:

$$d\rho_1 = \gamma_1 dt + {}_1 dW_t - K_1 \rho_1 (A + B)\, dt,$$
$$d\rho_2 = \gamma_2 dt + {}_2 dW_t - K_2 \rho_2 (A + B)\, dt.$$

Our original system now consists of one deterministic and two stochastic differential equations with constant diffusion coefficients, where W_t is the usual Wiener process (Brownian motion). These can be solved numerically using one of the standard SDE integrators such as Euler–Maruyama or the higher order Milstein method. For the simulations to follow, we used the Euler–Maruyama method with an extra modification. If the magnitude of either stochastic criminal generation rate δ_i is such that it would drive either ρ_i to a value smaller than 10% of that at steady state, then this parameter is set to zero for that specific step. For all simulations run, the noise strengths δ_i were equal to $\gamma_i/10$. In plots (a–d) of Figure 3, we present sample output from the stochastic simulations for the parameter values used in the deterministic case (left column). We also

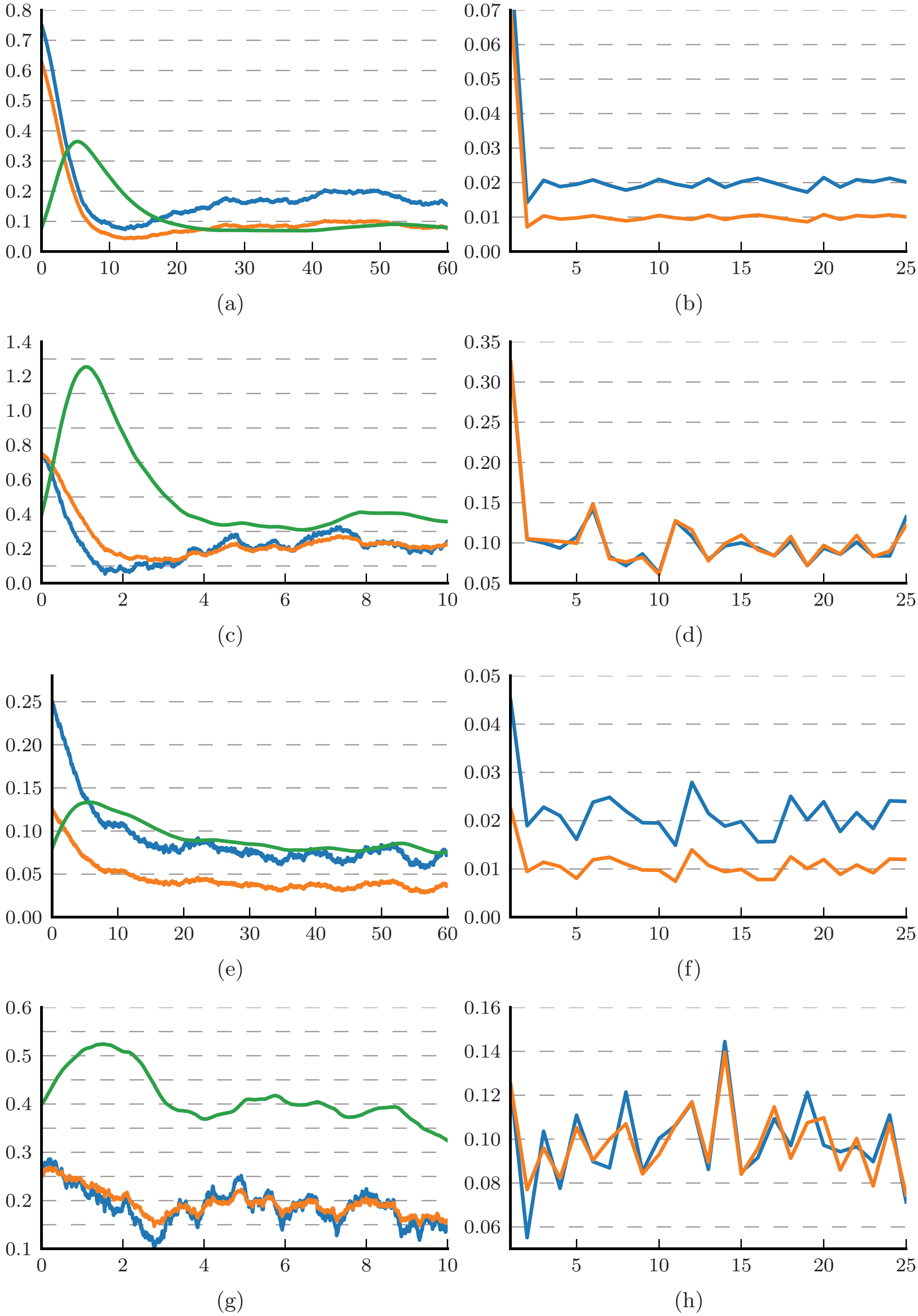

FIGURE 3. Numerical simulations of stochastic model using parameters described in Table 3. In the left column, we plot ρ_1 (red), ρ_2 (blue) and B (green), where the x-axis represents time in days. In the right column, we plot the monthly crime rate and the x-axis represents time in months.

Table 4. *Correlation coefficients (left columns) and their respective p-values (right columns) calculated from the output of the stochastic simulations shown in Figure 3*

	Pearson's r											
Plot	$S - \Delta S$		$M - \Delta M$		$S - M$		$\Delta S - \Delta M$		$S - \Delta M$		$M - \Delta S$	
a,b	−0.25	<0.01	−0.28	<0.01	1.00	<0.01	0.99	0.00	−0.30	<0.01	−0.23	<0.01
c,d	−0.55	<0.01	−0.48	<0.01	0.89	<0.01	0.80	<0.01	−0.24	<0.01	−0.57	<0.01
e,f	−0.38	<0.01	−0.42	<0.01	0.99	<0.01	0.99	0.00	−0.42	<0.01	−0.37	<0.01
g,h	−0.59	<0.01	−0.49	<0.01	0.89	<0.01	0.79	<0.01	−0.26	<0.01	−0.58	<0.01

present plots of the total number of crimes per month as predicted by the model (right column). This is achieved by summing

$$k_i \rho_i (A + B), \quad i = 1, 2$$

over 30 days. Calculating Pearson's r for the same categories as in the data analysis, we observe that agreement with the correlation coefficients obtained from the data analysis is in general satisfactory, as seen in Table 4.

It is evident from the simulations that the behaviour of the stochastic system is close to that of the deterministic, as expected. Altering $_1$ and $_2$ will change the amount of stochasticity present in the system and can be possibly used to improve agreement with data. We also note that the behaviour of both crime rates per month is qualitatively very similar to that of the actual time series, a promising fact for the ability of the model to approximate the data. It should be noted though that if any numerical comparison is to be made, a realistic set of parameter values is needed.

5 Parameter estimation

An inherent problem in most types of social models is that of estimating the value of the parameters used, coupled with the quantification of the uncertainty regarding their knowledge. When faced with such a problem, one would usually employ techniques such as non-dimensionalisation and experimentation to reduce the complexity of the model and obtain an estimate for the values of the parameters. Unfortunately, in many cases, this procedure may not be especially helpful or it may even be impossible to follow at all. For example, in this model, while non-dimensionalisation significantly reduces the number of parameters from 8 to 4, it fails to help with the estimation of parameters as we have no information regarding the baseline values of the constants that were used for the non-dimensionalisation (e.g. typical time scale). Due to the nature of the model, experimentation is not a valid approach either. Therefore, to obtain a realistic parameter regime, it is important to make full use of the crime data at our disposal.

We begin by choosing perhaps the simplest procedure that yields an estimate for the parameter values, minimising an objective function. In general, inverse problems associated with minimising functions can be hard to solve consistently as they are usually not well-posed. Instability of solutions is usually the culprit and this is something that

Table 5. *Parameter estimation for 29 months, each month being represented by a data point for every crime category*

k_1	k_2	λ_1	λ_2	γ_1	γ_2	ω	c_1	c_2	$\rho_1(0)$	$\rho_2(0)$	$B(0)$	A	$\mathscr{F}_{\min}$	Solver
0.50	1.05	0.30	0.08	0.81	0.77	0.03	1.05	1.23	0.91	0.33	0.53	1.40	0.06	TNC
0.69	1.87	0.54	0.69	0.35	0.66	0.01	0.46	1.05	0.38	0.11	2.17	1.18	0.07	BFGS
0.34	0.87	0.11	0.69	0.92	2.18	0.02	1.19	3.45	1.00	0.31	1.55	0.98	0.07	SLSQP

must be noted when performing this kind of parameter estimation, as it may significantly affect uncertainty regarding the estimates obtained. Keeping this in mind, we define an objective function $\mathscr{F}$, the arguments of which are the parameters whose value we want to estimate. This function will quantify the difference between the total monthly crime rates from the data analysis and that predicted by our model. One simple form for our function is the following:

$$\begin{aligned}\mathscr{F} = &\sum_{\text{All months}} (\text{Serious crime}_{\text{data}} - \text{Serious crime}_{\text{model}})^2 \\ &+ \sum_{\text{All months}} (\text{Minor crime}_{\text{data}} - \text{Minor crime}_{\text{model}})^2 .\end{aligned}$$

While executing the parameter estimation, the objective function is evaluated as follows:

(1) Fix initial data $\rho_1(0)$, $\rho_2(0)$, $B(0)$, and an initial guess for the vector of parameters to be estimated, defined as $p^0 = (k_1^0, k_2^0, \lambda_1^0, \lambda_2^0, \gamma_1^0, \gamma_2^0, \omega^0, c_1^0, c_2^0, \rho_1(0), \rho_2(0), B(0))$.
(2) Solve the deterministic model numerically.
(3) Evaluate Serious crime$_{\text{data}}$ (resp. Minor crime$_{\text{data}}$) by calculating $k_1\rho_1(A + B)$ (resp. $k_2\rho_2(A + B)$).
(4) Substitute into the objective function and calculate its value.
(5) Solve the minimisation problem, coupled with the constraint that all parameters must be greater than zero.

The minimisation was carried out using a plethora of solvers, such as Sequential Least Squares Quadratic Programming, Truncated Newton Constrained and the low-memory version of the BFGS algorithm (which is normally used for unconstrained optimisation, but can handle box constraints). A general observation is that the minimisation of $\mathscr{F}$ is not easy as the minimisers obtained may vary significantly depending on the initial guess as well as the algorithm chosen. One possible explanation would be the existence of lots or hills and/or valleys in the graph of $\mathscr{F}$, where most Newton-based algorithms struggle. Due to the nature of the problem, experimentation with the step size does not help either. Despite these difficulties in the minimisation procedure, it is possible to obtain estimates for the values of the parameters. Using these estimates (Table 5), we can solve the equations numerically and compare the actual crime rates with the model-predicted ones.

In Figure 4, we present example output from the simulations of the model. We performed two minimisations, first keeping the intrinsic attractiveness A constant (left column), and

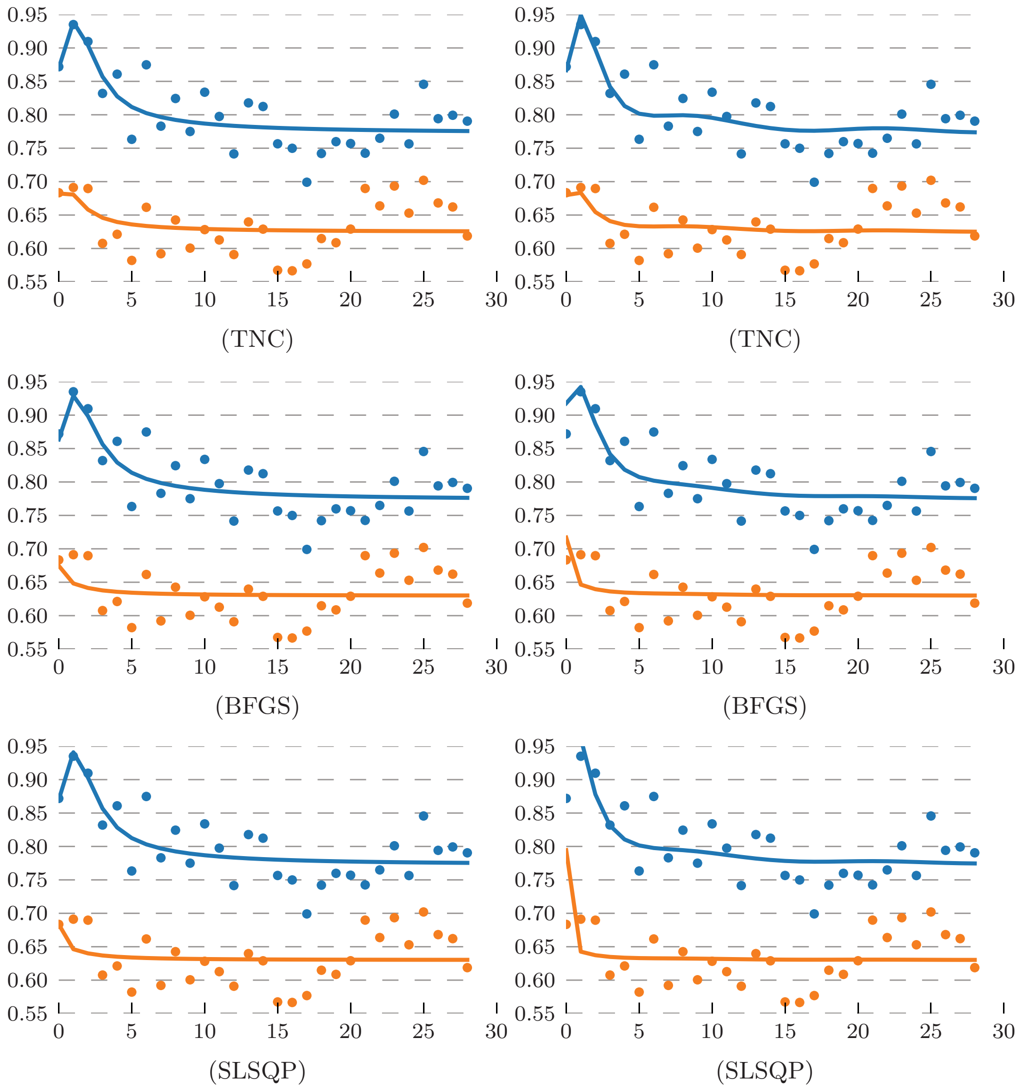

FIGURE 4. Comparison between actual crime data (dots) and deterministic model simulation (solid lines) for the Manchester area, using parameter values obtained from the minimisation procedure, as shown in Table 5. Plots in the left column were created with A constant, whereas its periodic form was used for plots in the right column. The x-axis represents time in months. Red colour represents serious crime and blue represents minor crime.

then using the parameter values obtained to estimate parameters for a more general periodic form of A with a period of one year, namely $A_1 + A_2 \sin\left(\pi t/6 + A_3\right)$ (right column). We observe that the deterministic model is a satisfactory approximation to the (random) data. However, due to the number of parameters estimated (either nine for constant A or 12 for periodic A), coupled with three initial conditions, the issue of possible parameter redundancy has to be considered. Indeed, it is possible to change the total crime rates without altering the equations of the model. To illustrate this, let us arbitrarily choose two positive numbers a_1, a_2 and consider the following scaling:

$$\rho_i \to a_i\rho_i, \quad k_i \to k_i/a_i, \quad \gamma_i \to a_i\gamma_i, \quad c_i \to a_ic_i, \quad \lambda_i \to \lambda_i/a_i, \quad i = 1,2.$$

It is now easy to see that all equations of the model and the total crime rates $k_i\rho_i(A+B)$ remain the same. This scaling shows that there cannot possibly be a set of unique parameters that serves as a global best fit to the crime rate, as we can always alter parameters by a_1, a_2 but leave everything else unchanged.

The availability of larger datasets would be a significant asset in improving model training, as would be a dataset of higher resolution (e.g. crimes reported daily instead of monthly). Furthermore, we performed the procedure above for areas other than Manchester, obtaining qualitatively similar results in the output of the simulations. The value of the objective function at its minimum was always consistent with the one is Manchester's case, ranging from 0.078 (Liverpool) to 0.045 (Southampton).

Finally, we can use the parameters obtained above for the stochastic model as well. We used the same parameters as for the deterministic simulations and added Gaussian noise. After running a number of simulations for each parameter set and averaging them, we present example output in Figure 5. We expect that, for a large enough number of simulations, the outcome of the stochastic model should be approximately the same as the deterministic one. Indeed, it is obvious that the stochastic simulations are qualitatively similar to the deterministic ones, as both crime rates fluctuate above and below their equilibrium values. A direct comparison however between the deterministic and the stochastic model is harder to make, as it is not easy to quantify their performance. We believe that the deterministic model should be considered as a tool for indicating the mean behaviour of both crime rates in the long run, whereas the output of the stochastic one indicates the expected variation in serious and minor crime and is qualitatively closer to the data.

6 Conclusions

As far as the actual data are concerned, it is clear that a variety of interesting conclusions can be drawn even from a preliminary analysis, as was the one conducted here. We based our model mainly on the high degrees of correlation between serious crime, minor crime and their rates of change. However, we chose not to incorporate in our model the fact that serious crime is correlated with the rate of change of minor crime, while there is no such link in the opposite direction. This observation is consistent with some theories proposed by criminologists [12, 17] and is something that needs to be further examined in the future, as it could simplify the model (possibly by imposing a relation between λ_1 and λ_2).

The transition from a deterministic model to a stochastic one was beneficial as it increased the flexibility of the model. One disadvantage, though, of this transition is the higher difficulty associated with estimating parameter values for a stochastic model. This can be sidestepped by using both versions of the model, estimating parameters from the deterministic one and then calibrating the amount of randomness present in the stochastic one to improve agreement with the data.

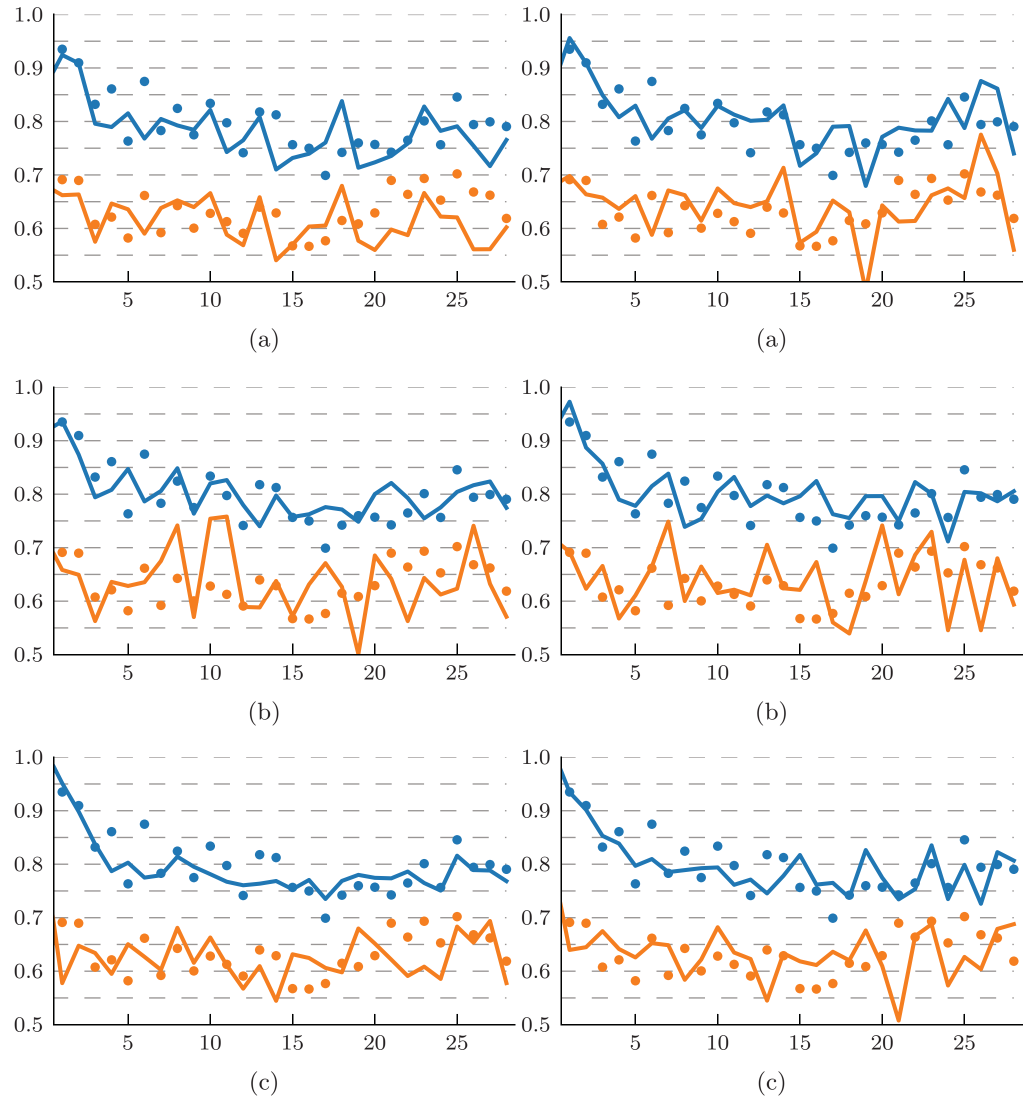

FIGURE 5. Comparison between actual crime data (dots) and stochastic model simulation (solid lines) for the Manchester area, using the same parameter values as in Figure 4 with $\delta_i = \gamma_i/10$ and A periodic. Red colour represents serious crime and blue represents minor. For the plots in the left column, 10 runs were executed and then averaged, as opposed to 100 runs in the right column.

The estimation of parameter values itself is also an area of further improvement. A clear cut solution to this problem may not be so easy to find although it is certain that more detailed data, especially in the temporal scale, would be an important asset. An objective function that more accurately represents the difference between actual and predicted crime of both types per month could theoretically improve the consistency of the minimisers obtained. It is also possible to adopt a different approach altogether in estimating parameter values, such as Bayesian methods, which is work in progress.

Finally, it should be mentioned that the framework set in this paper can be expanded to model crime distribution in space for serious and minor crime, along the lines of [18,20,21]. A model that successfully captures the points mentioned here and expands them to account for spatial variation as well is also work in progress. This could not only provide researchers with significantly improved understanding of the dynamics between serious and minor crime but also enable policy makers and law enforcement agencies to better predict crime evolution and increase their effectiveness as well as the allocation of their resources.

Acknowledgements

We wish to thank Prof. Shane Johnson of UCL and Dr Kevin Painter of Heriot-Watt University for many helpful discussions, as well as two anonymous reviewers for their valuable comments.

References

[1] Andresen, M. A. & Malleson, N. (2013) Crime seasonality and its variations across space. *Appl. Geography* **43**, 25–35.

[2] Blumstein, A. (2002) Crime modeling. *Oper. Res.* **50**(1), 16–24.

[3] Brantingham, P. J. & Brantingham, P. L. (1984) *Patterns in Crime*, Macmillan/McGraw-Hill School Division, New York.

[4] *Criminal activity data* (July 2014) URL: http://data.police.uk, accessed 2 February 2015.

[5] Epstein, J. M. (2002) Modeling civil violence: An agent-based computational approach. *Proc. Natl Acad Sci. USA* **99**(Suppl 3), 7243–7250.

[6] Evans, D., Fyfe, N. & Herbert, D. (2002) *Crime, Policing and Place: Essays in Environmental Criminology*, Taylor & Francis, London.

[7] Gordon, M. B., Iglesias, J. R., Semeshenko, V. & Nadal, J. P. (2009) Crime and punishment: The economic burden of impunity. *Eur. Phys. J. B* **68**(1), 133–144.

[8] Gorr, W., Olligschlaeger, A. & Thompson, Y. (2003) Short-term forecasting of crime. *Int. J. Forecast.* **19**(4), 579–594.

[9] Johnson, S. D., Bowers, K. J. & Hirschfield, A. (1997) New insights into the spatial and temporal distribution of repeat victimization. *Br. J. Criminology* **37**(2), 224–241.

[10] Johnson, S. D. & Bowers, K. J. (2004) The burglary as clue to the future the beginnings of prospective hot-spotting. *Eur. J. Criminology* **1**(2), 237–255.

[11] Kovandzic, T. V. & Sloan, J. J. (2002) Police levels and crime rates revisited: A county-level analysis from Florida (1980–1998). *J. Criminal Justice* **30**(1), 65–76.

[12] Lahey, B. B., Moffitt, T. E. & Caspi, A. (2003) *Causes of Conduct Disorder and Juvenile Delinquency*, Guilford Press, New York.

[13] Malleson, N., Heppenstall, A. & See, L. (2012) Using an agent-based crime simulation to predict the effects of urban regeneration on individual household burglary risk. *Environ. Plann. B Plan.* **40**(3), 405–426.

[14] McCalla, S. G., Short, M. B. & Brantingham, P. J. (2012) The effects of sacred value networks within an evolutionary, adversarial game. *J. Stat. Phys.* **151**(3), 1–16.

[15] Mohler, G. O., Short, M. B., Brantingham, P. J., Schoenberg, F. P. & Tita, G. E. (2011) Self-exciting point process modeling of crime. *J. Am. Stat. Assoc.* **106**(493), 100–108.

[16] Nuño, J. C., Herrero, M. A. & Primicerio, M. (2008) A triangle model of criminality. *Physica A: Stat. Mech. Appl.* **387**(12), 2926–2936.

[17] Sampson, R. J. & Raudenbush, S. W. (1999) Systematic social observation of public spaces: A new look at disorder in Urban neighborhoods. *Am. J. Soc.* **105**(3), 603–651.

[18] Short, M. B., Bertozzi, A. L. & Brantingham, P. J. (2010) Nonlinear patterns in Urban crime: Hotspots, bifurcations, and suppression. *SIAM J. Appl. Dyn. Syst.* **9**(2), 462–483.

[19] Short, M. B., Brantingham, P. J. & D'orsogna, M. R. (2010) Cooperation and punishment in an adversarial game: How defectors pave the way to a peaceful society. *Phys. Rev. E* **82**(6), 066114.

[20] Short, M. B., D'orsogna, M. R., Brantingham, P. J. & Tita, G. E. (2009) Measuring and modeling repeat and near-repeat burglary effects. *J. Quant. Criminology* **25**(3), 325–339.

[21] Short, M. B., D'orsogna, M. R., Pasour, V. B., Tita, G. E., Brantingham, P. J., Bertozzi, A. L. & Chayes, L. B. (2008) A statistical model of criminal behavior. *Math. Models Methods Appl. Sci.* **18**(supp01), 1249–1267.

[22] Short, M. B., Mohler, G. O., Brantingham, P. J. & Tita, G. E. (2012) Gang rivalry dynamics via coupled point process networks. *Discrete and Continuous Dynamical Systems*, **19**(5), 1459–1477.

[23] Sooknanan, J., Bhatt, B. & Comissiong, D. M. G. (2013) Catching a gang - a mathematical model of the spread of gangs in a population treated as an infectious disease. *Int. J. Pure Applied Math.* **83**(1), 25–43.

[24] Tompson, L., Johnson, S., Ashby, M., Perkins, C. & Edwards, P. (2014) UK open source crime data: Accuracy and possibilities for research. *Cartography Geogr. Inform. Sci.* **42**(2), 97–111.

Appendix A

Table A 1. *Pearson's r (top half) and Spearman's ρ (bottom half) calculated using data from September 2011 to January 2014 obtained from the UK police website. The coefficients are in the left columns whereas the p-values are on the right. S and M represent serious and minor crime respectively and ΔS and ΔM are forward differences, e.g.* ΔS_{June} = *(Total amount of serious crime in July) − (Total amount of serious crime in June)*

	Pearson's *r*											
	$S-\Delta S$		$M-\Delta M$		$S-M$		$\Delta S-\Delta M$		$S-\Delta M$		$M-\Delta S$	
Birmingham	−0.58	0.01	−0.48	0.02	0.67	<0.01	0.58	<0.01	−0.56	0.01	−0.09	0.69
Manchester	−0.52	<0.01	−0.52	<0.01	0.58	<0.01	0.80	0.00	−0.58	<0.01	−0.29	0.13
Leeds–Bradford	−0.47	0.01	−0.55	<0.01	0.58	<0.01	0.71	0.00	−0.46	0.01	−0.22	0.26
Liverpool–Merseyside	−0.58	<0.01	−0.66	<0.01	0.67	<0.01	0.71	0.00	−0.35	0.07	−0.55	<0.01
Southampton–Portsmouth	−0.33	0.09	−0.40	0.04	0.66	<0.01	0.44	0.02	−0.17	0.38	−0.22	0.26
Newcastle–Sunderland	−0.54	<0.01	−0.55	<0.01	0.61	<0.01	0.56	<0.01	−0.22	0.26	−0.38	0.04
Nottingham–Derby	−0.53	<0.01	−0.57	<0.01	0.57	<0.01	0.45	0.02	−0.51	0.01	0.01	0.98
Sheffield	−0.57	<0.01	−0.51	0.01	0.44	0.02	0.49	0.01	−0.51	0.01	−0.08	0.67
Suffolk	−0.32	0.09	−0.48	0.01	0.85	0.00	0.54	<0.01	−0.17	0.38	−0.25	0.21
Norfolk	−0.57	<0.01	−0.52	<0.01	0.54	<0.01	0.21	0.29	−0.11	0.57	−0.15	0.46
Cambridgeshire	−0.42	0.03	−0.50	0.01	0.52	<0.01	0.53	<0.01	−0.24	0.22	−0.24	0.22
Lincolnshire	−0.49	0.01	−0.51	0.01	0.73	0.00	0.76	0.00	−0.50	0.01	−0.26	0.18
Cumbria	−0.51	0.01	−0.51	0.01	0.68	<0.01	0.61	<0.01	−0.38	0.04	−0.28	0.15
Durham	−0.45	0.02	−0.50	0.01	0.56	<0.01	0.44	0.02	−0.36	0.06	−0.01	0.95
Sussex	−0.69	<0.01	−0.45	0.02	0.77	0.00	0.78	0.00	−0.59	<0.01	−0.33	0.08
Devon–Cornwall	−0.53	<0.01	−0.40	0.04	0.66	<0.01	0.71	0.00	−0.47	0.01	−0.21	0.29
Entire dataset	−0.61	<0.01	−0.60	<0.01	0.69	0.00	0.87	0.00	−0.65	<0.01	−0.43	0.02
Birmingham	−0.47	0.03	−0.58	<0.01	0.66	<0.01	0.57	0.01	−0.63	<0.01	−0.09	0.70
Manchester	−0.54	<0.01	−0.53	<0.01	0.55	<0.01	0.77	0.00	−0.60	<0.01	−0.30	0.12
Leeds–Bradford	−0.46	0.01	−0.59	<0.01	0.62	<0.01	0.65	<0.01	−0.51	0.01	−0.24	0.23
Liverpool–Merseyside	−0.64	<0.01	−0.60	<0.01	0.69	0.00	0.75	0.00	−0.41	0.03	−0.46	0.01
Southampton–Portsmouth	−0.47	0.01	−0.39	0.04	0.62	<0.01	0.55	<0.01	−0.14	0.47	−0.26	0.18
Newcastle–Sunderland	−0.57	<0.01	−0.56	<0.01	0.71	0.00	0.38	0.04	−0.19	0.33	−0.46	0.01
Nottingham–Derby	−0.55	<0.01	−0.54	<0.01	0.45	0.02	0.40	0.03	−0.53	<0.01	0.02	0.94
Sheffield	−0.58	<0.01	−0.48	0.01	0.40	0.03	0.55	<0.01	−0.52	<0.01	−0.02	0.92
Suffolk	−0.24	0.22	−0.51	0.01	0.84	0.00	0.46	0.01	−0.18	0.35	−0.24	0.22
Norfolk	−0.65	<0.01	−0.44	0.02	0.43	0.02	0.18	0.36	−0.20	0.30	−0.08	0.70
Cambridgeshire	−0.47	0.01	−0.48	0.01	0.49	0.01	0.58	<0.01	−0.21	0.29	−0.30	0.13
Lincolnshire	−0.44	0.02	−0.49	0.01	0.73	0.00	0.71	0.00	−0.50	0.01	−0.29	0.14
Cumbria	−0.47	0.01	−0.51	0.01	0.62	<0.01	0.55	<0.01	−0.34	0.08	−0.26	0.19
Durham	−0.32	0.09	−0.49	0.01	0.54	<0.01	0.40	0.03	−0.28	0.15	<0.01	0.99
Sussex	−0.71	0.00	−0.36	0.06	0.80	0.00	0.75	0.00	−0.56	<0.01	−0.40	0.04
Devon–Cornwall	−0.47	0.01	−0.35	0.07	0.69	0.00	0.71	0.00	−0.39	0.04	−0.16	0.43
Entire dataset	−0.59	<0.01	−0.54	<0.01	0.64	<0.01	0.90	0.00	−0.64	<0.01	−0.41	0.03